Accompanying Documents

1. Marked-up copies incorporating the amendments made herein (Appendix A).
2. Clean copies of the pending claims after incorporation of the amendments made herein (Appendix B).
3. Declaration of the Inventors.
4. Printout of e-mails from GenBank and "The Plant Cell."

AMENDMENT

In the Specification:

Please replace the paragraph in the specification at page 11, line 31, to page 12, line 13, with the following:

Sub > A1 B1

Figure 3 depicts alignment of cytochrome P450 proteins that exhibited the most similarity to DWF4 (Seq ID NO.:2) in BLAST searches. GenBank accession numbers are AF044216 (DWF4; CYP90B)(Seq ID NO.:2), X87368 (CPD; CYP90A), U54770 (tomato; CYP85), D64003 (cyanobacteria; CYP120), U32579 (maize; CYP88), U68234 (zebrafish; CYP26), and M13785 (human; CYP3A3X). Dashes indicate gaps introduced to maximize alignment. Domains indicated in Figure 2B are highlighted in a box. Amino acid residues that are conserved >50% between the compared sequences are highlighted by a reverse font, and identical residues between DWF4 and CPD are boxed and italicized. Open triangles are placed under the 100% conserved residues. Closed triangles locate functionally important amino acid residues, for example, threonine (T) at 369, which is thought to bind molecular oxygen, and cysteine (C) at 516, which links to a heme prosthetic group by a thiolate bond. X's indicate mutated residues in *dwf4* alleles. Multiple sequence alignment was performed using PILEUP in the Genetics Computer Group package,

A' cont'd

and box shading was made possible by the ALSCRIPT package (Barton (1993) *Protein Eng*. 6:37-40). --

Please replace the paragraph in the specification at page 17, line 1 to page 18, line 5, with the following.

A2

--Techniques for determining nucleic acid and amino acid "sequence identity" are known in the art. Typically, such techniques include determining the nucleotide sequence of the mRNA for a gene and/or determining the amino acid sequence encoded thereby, and comparing these sequences to a second nucleotide or amino acid sequence. In general, "identity" refers to an exact nucleotide-to-nucleotide or amino acid-to-amino acid correspondence of two polynucleotides or polypeptide sequences, respectively. Two or more sequences (polynucleotide or amino acid) can be compared by determining their "percent identity." The percent identity of two sequences, whether nucleic acid or amino acid sequences, is the number of exact matches between two aligned sequences divided by the length of the shorter sequences and multiplied by 100. An approximate alignment for nucleic acid sequences is provided by the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2:482-489 (1981). This algorithm can be applied to amino acid sequences by using the scoring matrix developed by Dayhoff, Atlas of Protein Sequences and Structure, M.O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA, and normalized by Gribskov, Nucl. Acids Res. 14(6):6745-6763 (1986). An exemplary implementation of this algorithm to determine percent identity of a sequence is provided by the Genetics Computer Group (Madison, WI) in the "BestFit" utility application. The default parameters for this method are described in the Wisconsin Sequence Analysis Package Program Manual, Version 8 (1995) (available from Genetics Computer Group, Madison, WI). A preferred method of establishing percent identity in the context of the present invention is to use the MPSRCH package of programs copyrighted by the University of Edinburgh, developed by John F. Collins and Shane

A2 cont'd

S. Sturrok, and distributed by IntelliGenetics, Inc. (Mountain View, CA). From this suite of packages the Smith-Waterman algorithm can be employed where default parameters are used for the scoring table (for example, gap open penalty of 12, gap extension penalty of one, and a gap of six). From the data generated the "Match" value reflects "sequence identity." Other suitable programs for calculating the percent identity or similarity between sequences are generally known in the art, for example, another alignment program is BLAST, used with default parameters. For example, BLASTN and BLASTP can be used using the following default parameters: genetic code = standard; filter = none; strand = both; cutoff = 60; expect = 10; Matrix = BLOSUM62; Descriptions = 50 sequences; sort by = HIGH SCORE; Databases = non-redundant, GenBank + EMBL + DDBJ + PDB + GenBank CDS translations + Swiss protein + Spupdate + PIR.

Please replace the paragraph in the specification at page 51, lines 25-30, with the following:

A3

--Analysis of the complete genomic sequence, starting at the EcoRI site, with the promoter prediction by neural network (NNPP) package, indicated that the gene included a putative promoter (TATAT is found in the putative promoter region between nucleotides -143 to -78) and polyadenylation signal sequences (AATAA near a position at 3238 bp and a putative GU-rich signature from 3283 to 3290 bp).--

Please replace the paragraph in the specification at page 55, line 28, to page 56, line 13, with the following:

~~A4~~ sub C3

--Thus, phylogenetic analyses of these seven proteins with cytochrome P450s unique to plants (group A; Durst and Nelson (1995), *supra*) indicate that DWF4 does not cluster with these cytochrome P450s (Figure 4). Rather, DWF4 clustered with cytochrome P450s from other organisms: cyanobacteria (CYP120), rat (CYP3A2), human (CYP3A3X), and plants (CYP90, CYP85, and CYP88). DWF4 also deviates from the consensus sequence in the group A heme binding domain in that it